Ferris, B.G., Dockery, D.W., Ware, J.H., Speizer, F.E., Spiro, R.
"The Six-City Study: Examples of Problems in Analysis of the Data"
Environmental Health Perspectives 52: 115-123, 1983.

ABSTRACT. This paper presents some of the results from cross-sectional analyses and studies during air pollution alerts obtained as a part of the Six-City Study, a longitudinal study of the respiratory effects of air pollution. These analyses illustrate some of the limitations and uncertainties of epidemiologic studies. For example, an earlier report noted increased respiratory illness rates for children living in homes where gas was used for cooking. A later analysis did not confirm this. Reasons for this are explored by using different criteria and variables to be controlled for. The results illustrate that the strength of the association between cooking fuel and illness was sensitive to the definitions of the variables and the number of subjects and city cohorts. Similar examples are presented for illness rates for four respiratory diseases: asthma, bronchitis, illness before age 2 and illness last winter. These examples of cross-sectional analyses emphasize the ambiguities of studies of possible health effects of air pollution exposures close to the present ambient air quality standards.

Environmental Health Perspectives
Vol. 52, pp. 115-123, 1983

The Six-City Study: Examples of Problems in Analysis of the Data

by B. G. Ferris, Jr.,* D. W. Dockery,* J. H. Ware,† F. E. Speizer‡ and R. Spiro III*

This paper presents some of the results from cross-sectional analyses and studies during air pollution alerts obtained as a part of the Six-City Study, a longitudinal study of the respiratory effects of air pollution. These analyses illustrate some of the limitations and uncertainties of epidemiologic studies. For example, an earlier report noted increased respiratory illness rates for children living in homes where gas was used for cooking. A later analysis did not confirm this. Reasons for this are explored by using different criteria and variables to be controlled for. The results illustrate that the strength of the association between cooking fuel and illness was sensitive to the definitions of the variables and the number of subjects and city cohorts. Similar examples are presented for illness rates for four respiratory diseases: asthma, bronchitis, illness before age 2 and illness last winter. These examples of cross-sectional analyses emphasize the ambiguities of studies of possible health effects of air pollution exposures close to the present ambient air quality standards.

Introduction

This paper describes some results from cross-sectional analyses and studies during air pollution alerts performed as part of the Six-City Study, a longitudinal study of the respiratory health effects of air pollution. These analyses illustrate some of the limitations and uncertainties of epidemiologic studies. We begin with a brief outline of the study design.

The Six-City Study was designed to test the adequacy of the present federal standards for SO_2 and particulates, to develop data on the effects of small particles, to assess the representativeness of a central station as an index of exposure and to assess the effect of the home environment in modifying exposure as indicated by outdoor levels. We planned to study primarily chronic effects but it was soon apparent that short-term fluctuations in air pollutants, particularly in Steubenville, offered the chance to study possible acute effects.

The six cities studied were selected on the basis of their historical levels of pollutants to include clean cities, cities close to the present standards for SO_2 and particulates, and cities above the standards. The six cities selected are listed in Table 1. The time of year that the city was visited also is indicated.

Methods

Random samples of adults were selected for study in each city from various census lists that were available. We attempted to obtain 1500 adults per city. The children selected were first and second graders. Additional cohorts of chil-

Table 1. Cities studied by pollution category and time of year.

City	Season	Pollution category
Portage, WI	Fall	Clean
Topeka, KA	Spring	
Watertown, MA	Fall	Somewhat below standard
Kingston-Harriman, TN	Spring	
St. Louis, MO	Fall	Above standard
Steubenville, OH	Spring	

*Department of Physiology, Harvard School of Public Health, Boston, MA 02115.
†Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115.
‡Channing Laboratory, Department of Medicine, Harvard Medical School and Brigham and Women's Hospital and Department of Physiology, Harvard School of Public Health, Boston, MA 02115.

dren were added in subsequent years from the new first grade and new additions to the grades under study until about 2000 children were enrolled in each city. The adults were studied every third year and the children every year. Contact was maintained with the adult sample by mailing annual reports of the study. Respiratory health effects in both adults and children were assessed by means of a standard respiratory questionnaire and simple tests of pulmonary function. Smoking and occupational histories were obtained on the adults. Smoking habits were asked of the children when they reached the fourth grade or at age 9 years. Home characteristics such as air conditioning, heating and cooking fuels and number of smokers in the home also were tabulated.

Air monitoring was carried out in each city at a central site and at satellite sites in the community. Indoor and outdoor levels of a variety of pollutants were measured at each satellite site. Personal sampling was carried out to assess whether our modeling of estimated exposure has been reasonable. The pollutants and sampling methods are summarized in Table 2. The continuous methods are calibrated by our own team as well as by the Environmental Protection Agency. EPA also makes quality checks on our analytic procedures at regular intervals. The collection of bubbler samples is organized to coincide with the regular EPA sampling schedule. A more detailed description of the study is given in a previous report (*1*).

Results

Let us now look at some of the data and analyses and examine some of the problems we see in these analyses.

Steubenville Alert Study

In an attempt to quantify the impact of acute air pollution exposures on pulmonary function, we measured the pulmonary function of approximately 200 school children in Steubenville, Ohio, before, during and following periods of high air pollution concentrations.

In the fall of 1978, baseline measurements of pulmonary function were obtained on third and fourth graders between October 16 and 23. On November 5, total suspended particulate (TSP) concentrations reached 422 µg/m³ at the offices of the Northern Ohio Valley Air Authority (NOVAA) in Steubenville (Fig. 1). Because of previously determined criteria involving precipitation forecasting and levels of air pollution, NOVAA declared an "alert." The 24-hr mean sulfur dioxide (SO_2) concentration was 211 µg/m³ on November 5 and climbed to 272 µg/m³ on November 6. The National Primary Ambient Air Quality Standard for 24-hr concentrations of TSP is 260 µg/m³ and for SO_2 is 365 µg/m³. The children were retested on November 6 and 7 and weekly for the next 3 weeks. Each child's pulmonary

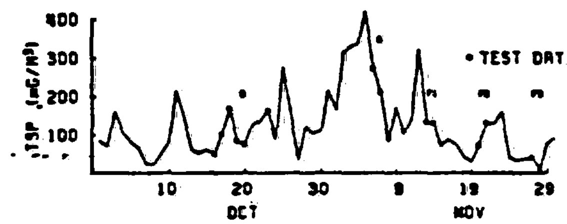

FIGURE 1. Fall 1978 alert. TSP concentrations measured at the NOVAA offices in Stuebenville, OH. Spirometric testing days are indicated by squares. (B) = baseline, (A) = alert, (F1) = follow-up 1, (F2) = follow-up 2 and (F3) = follow-up 3.

Table 2.

	Pollutant	Period	Sampler	Additional analysis
Central site	SO_2, NO_2, O_3	1 hr	Continuous	
	SO_2, NO_2	24 hr	Bubbler	
	TSP	24 hr	Hi-volume	Sulfates
	Fine and coarse particles[a]	24 hr	Dichotomous	X-Ray fluorescence
	Mass respirable particles[b]	24 hr	Cyclone	Neutron activation Sulfates
Satellite (indoor and outdoor)	SO_2, NO_2	24 hr	Bubbler	
	Mass respirable particles[b]	24 hr	Cyclone	Neutron activation Sulfates
Personal	Mass respirable particles[b]	24 hr	Cyclone	Sulfates
	NO_2	7 day	Palmes tubes	

[a]"Fine" defined as < 2.5 µm aerodynamic diameter; "coarse" defined as 2.5-15 µm aerodynamic diameter.
[b]50% cut at 3.5 µm aerodynamic diameter.

function measurements during the alert (A) and at follow-up tests (F1, F2, F3) were compared to their own values during the baseline tests (B). Thus, sex and race are not considered in the analysis.

For those children seen on all five visits, $FEV_{1.0}$ averaged 12 mL lower at the alert and follow-up visits compared to the baseline values (Fig. 2). The mean $FEV_{1.0}$ at the baseline measurement of these children was 1.686 L. At the first follow-up visit (F1), mean decline in $FEV_{1.0}$ was 11 mL; in the second follow-up it was 15 mL; and in the third follow-up it was only 5 mL. As suggested by the standard errors about the mean, none of the declines reached a statistically significant difference from zero at the 0.05 level. The trend in the differences is consistent with a decline in the $FEV_{1.0}$ following the air pollution alert which may last several weeks. Other factors might have prevailed. We have examined the effect of levels of pollution on the day of the measurement and have explored the possibility that a few children could be driving the overall effect seen. We have also attempted to assess the effects of boredom of the children or the technicians.

If we subtract from each $FEV_{1.0}$ measurement the expected value for that child based upon a prediction formula using sex, race and height, we can correlate the residual $FEV_{1.0}$ values with the daily pollution concentrations for the 24 hr ending at 8 A.M. of the morning of the pulmonary function testing. This analysis is sensitive to a very short-term effect, with recovery within 24 hr. As seen in Figure 3, there is little correlation between mean residual $FEV_{1.0}$ and TSP concentrations that day for this study. The correlations with the other pollutants, SO_2 and NO_2, were not as high as that for TSP.

The study was repeated in the fall of 1979. Initially, the intent was to examine the effect of repeated weekly measurements of pulmonary function without an air pollution episode. Baseline measurements were made in October. On November 16, a "sham" alert was declared, and the children were restudied.

On November 20, TSP concentrations reached 271 $\mu g/m^3$ and SO_2 concentrations reached 439 $\mu g/m^3$. Both values exceeded the 24-hr standard. An alert was called. Children in school on November 21, which was the day before the start of Thanksgiving vacation, were retested. All the children were retested weekly for the following 3 weeks. Differences in $FEV_{1.0}$ compared to baseline (Fig. 4) showed an increase on the sham alert day (S) and a decline following the actual alert. Although larger than the measured declines in 1978, the differences were not significantly different from zero at the 0.05 level for those children who were tested all six times. As in the 1978 alert, the $FEV_{1.0}$ residuals did not correlate well with the daily TSP, SO_2 or NO_2 concentrations.

The following spring, in 1980, these same children were tested again in a sham alert for five consecutive weeks. $FEV_{1.0}$ declined in a pattern very similar in shape and magnitude to that seen in the previous alert studies (Fig. 5). The air

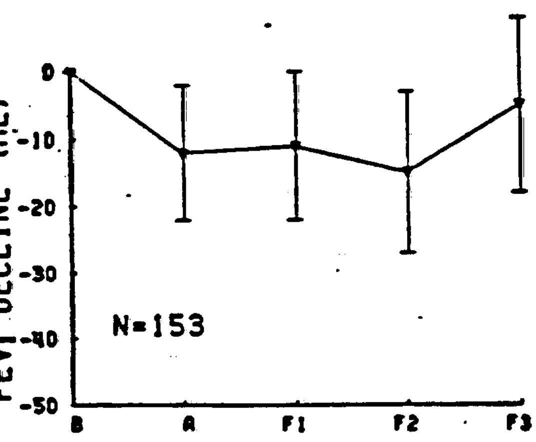

FIGURE 2. Mean difference in $FEV_{1.0}$ compared to baseline measurement for those children tested on all five visits in the fall 1978 alert study. Vertical bars denote one standard error above and below the mean.

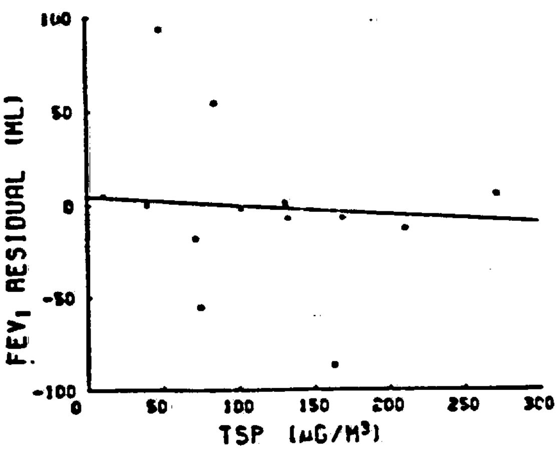

FIGURE 3. Mean residual $FEV_{1.0}$ for each testing day in the fall 1978 alert study plotted against mean TSP concentrations for the 24 hr ending at 8 A.M. for that day. The best-fit line, based on least squares weighted by the number of children tested each day, is also shown ($R = 0.15$).

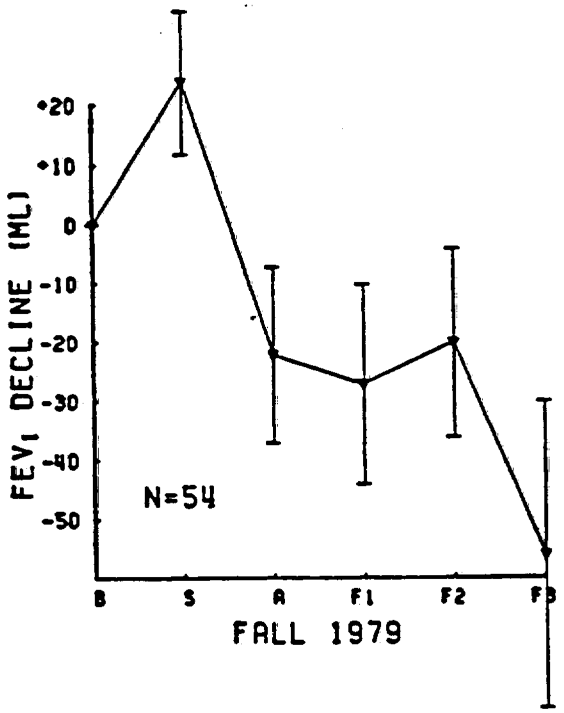

FIGURE 4. Mean difference in $FEV_{1.0}$ compared to baseline, with standard errors, for those children tested on all six visits during the fall 1979 alert study.

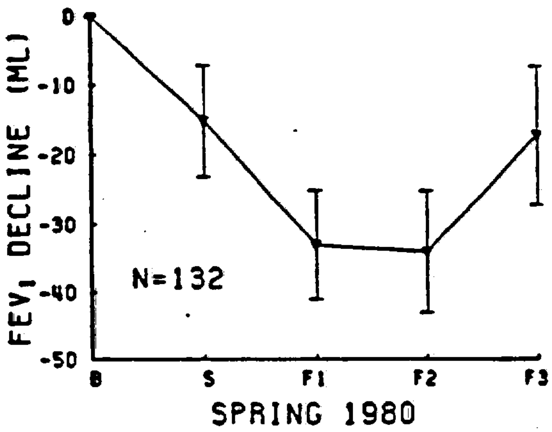

FIGURE 5. Mean difference in $FEV_{1.0}$ compared to baseline, with standard errors, for those children tested on all five visits during the spring 1980 sham alert.

pollution concentrations, as illustrated by the daily mean TSP level (Fig. 6), were relatively low. The maximum TSP was 220 μg/m³, and the maximum SO_2 was 169 μg/m³. This suggested that the observed declines might be a result of fatigue, or lack of interest by the children. On the other hand, there was a very high correlation between the mean residual $FEV_{1.0}$ for each day and the TSP concentration in the spring 1980 study (Fig. 7).

In a further attempt to understand the changes in pulmonary function which had been observed, these same children were tested again in the fall of 1980 for five consecutive weeks (Fig. 8). During this period, TSP had a maximum concentration of 159 μg/m³, and SO_2 a maximum concentration of 166 μg/m³. $FEV_{1.0}$ again showed declines similar to those seen following the air pollution alerts. The correlation of daily TSP with residual $FEV_{1.0}$ was fairly good ($R^2 = 0.204$).

Discussion

These four studies provide suggestive but inconclusive evidence concerning the relationship between short-term changes in air pollution concentration and the level of $FEV_{1.0}$. Taken together, these children exposed to sudden increases of TSP and SO_2 concentrations near or above the current 24-hr standards had declines of 1 to 2% in $FEV_{1.0}$. These changes are small relative to sampling variability and are marginally significant.

We also need to evaluate the medical significance of such changes in pulmonary function. That is, do they lead to a permanent effect? Initial cross-sectional comparisons of the levels of pulmonary function show that Steubenville children have values comparable to those in the other cities in spite of presumably having experienced similar high, or higher, levels of short-term TSP and SO_2 peaks in the past. To reduce the potentially confounding effects of other factors, we are following these children prospectively to compare lung function development in these children to that in similarly aged children from other cities in which no such short-term peak exposures have occurred.

Children's Illness Prevalence

The second set of analyses to be discussed describes associations between illness rates for children living in the six cities and potential risk factors such as passive smoking or the use of gas cooking fuel. The following points will be emphasized. First, the effects of environmental risk factors on respiratory disease rates are sensitive to what may be termed the 'style' of analysis adopted. That is, different results may be obtained depending on: (a) definitions of the risk factors; (b) the composition and selection of the sample analyzed; and (c) the set of additional

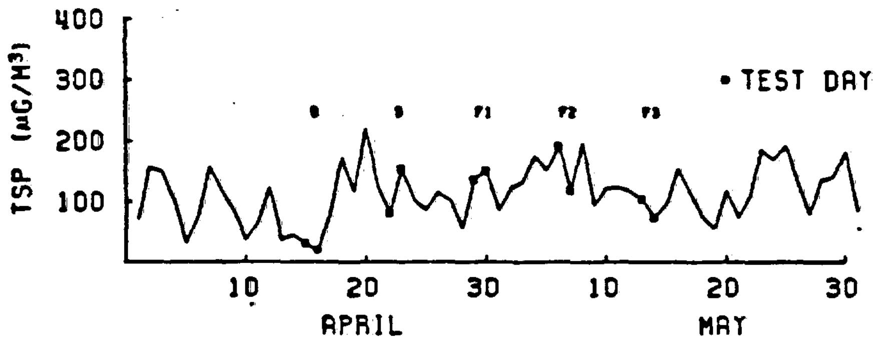

FIGURE 6. TSP concentrations for 24 hr measured in Steubenville during the spring 1980 sham alert with spirometric testing days indicated.

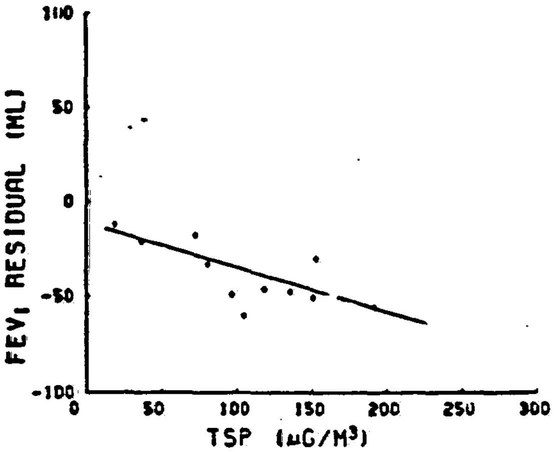

FIGURE 7. Mean residual $FEV_{1.0}$ for each testing day in the spring 1980 sham alert plotted against mean TSP concentrations for the 24 hr ending at 8 A.M. for that day.

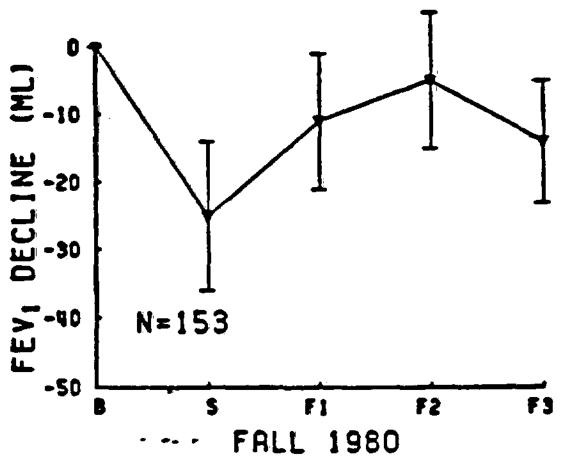

FIGURE 8. Mean difference in $FEV_{1.0}$ compared to baseline measurements and standard errors during the fall 1980 sham alert study.

variables included in the analysis. This sensitivity will be illustrated in the analyses of the prevalence of history of respiratory illness before age 2. Second, information will be presented on rates by city for four respiratory diseases: asthma, bronchitis, illness before age 2, and respiratory illness last year.

In a recent attempt to extend results reported earlier by our group (2) on the relationship between prevalence of respiratory illness before age 2 and cooking fuel, we repeated the analysis after including an additional four cohorts of children enrolled during 1978-1979. In this reanalysis, the association between illness before age 2 and cooking fuel was not statistically significant at the level of $p < 0.05$. This led to a closer comparison of the two analyses to determine why the result was not more closely replicated.

There was one obvious difference noted between the original and the more recent analysis. In the original analysis the data from 10 city-cohorts were pooled over sex. In the recent analysis, data from 14 city-cohorts (an additional 1493 children) were examined separately by sex. On closer examination, several other differences in variable definition were noted between the two analyses. In the first analysis, the cooking variable contrasted homes using gas only with those using electricity only. In the more recent analysis, the cooking variable contrasted homes using some gas with those using any other fuel. In the earlier analysis, the smoking variable indicated whether anyone in the home smoked. In the later analysis, this variable indicated only maternal smoking. Finally, socioeconomic status (SES) was defined differently in the two analyses. In the

first, it was based on both parents' education and occupation. In the second, it was based on a simple average of the parents' education.

Table 3 gives the strength of the association between illness before age 2 and cooking fuel in various analyses conducted to investigate the difference between the original result and the more recent analysis. Results are given for the pooled data and for boys and girls examined separately. For each analysis, the number of subjects is given, and the χ^2 test with one degree of freedom for association between cooking fuel and illness is presented. Finally, Table 3 presents the probability value for each of the analyses.

In analysis 1 in Table 3, reported earlier by us (2), the data from 5660 children in 10 city-cohorts were pooled over sex and analyzed. There was a significant relationship between illness and cooking fuel ($p = 0.01$). Children from homes with gas cooking had higher rates of illness before age 2 than did children from homes with electric cooking. When boys and girls were analyzed separately, the relationship between illness and cooking was significant for girls ($p = 0.02$), but not for boys ($p = 0.20$), though the direction of the association was the same in both groups.

Analysis 2 was our initial reanalysis with the four added cohorts, and newly defined variables for SES, cooking fuel, and household smoking. The association between cooking and illness before age 2 did not achieve statistical significance either for girls ($p = 0.12$) or for boys ($p = 0.48$). Because of this difference in the strength of the association, we believed it would be important to test individually now changes in the analysis affected the results.

Analysis 3 is comparable to analysis 1, except that SES and the cooking variable were defined as in analysis 2. That is, it contrasted homes using any gas cooking with homes using electricity or other cooking fuels. Formerly the cooking variable had contrasted homes using only gas with homes using only electricity. In addition the definition of SES used parental education only. These changes added 167 children to the data set compared to analysis 1. For the data pooled over sex, the relationship between illness and cooking was significant ($p = 0.04$). When the data were stratified by sex and analyzed separately, the relationship between illness and cooking was significant for girls ($p = 0.03$), but not for boys ($p = 0.49$).

In analysis 4 we added the four new cohorts. SES, smoking, and cooking were defined as in analysis 2. For the data pooled over sex, the relationship between illness and cooking was significant ($p = 0.04$). When the data were stratified by sex and analyzed separately, the relationship between illness and cooking was significant for girls ($p = 0.04$), but not for boys ($p = 0.44$).

To explain the discrepancy between analyses 2 and 4, we assessed the differences in the specific models used. In analysis 4, the effect of gas cooking was based on the original model used in analysis 1. It included an interaction term for city-cohort, smoking, and illness prevalence, but did not include an interaction term for SES and illness. In our new analysis (analysis 2) education and illness prevalence were significantly associated and there was no interaction between city-cohort, smoking, and illness. While the original model provided a satisfactory fit to the earlier data, it did not provide as adequate a fit to the enlarged data set as did a new model based on these data.

These results illustrate several problems in ex-

Table 3. Results of several analysis of the association between gas cooking and respiratory illness before age 2.

Analysis	N	χ^2	p
1 Original data, original model			
Pooled	5660	6.70	0.01
Boys	2936	1.59	0.20
Girls	2724	5.26	0.02
2 Enlarged data set, new model[a]			
Boys	3721	0.51	0.48
Girls	3433	2.44	0.12
3 Original data, SES, cooking fuel redefined			
Pooled	5827	4.16	0.04
Boys	3013	0.47	0.49
Girls	2814	4.61	0.03
4 Enlarged data set, SES, cooking fuel, smoking redefined			
Pooled	7153	4.13	0.04
Boys	3721	0.61	0.44
Girls	3432	4.06	0.04

[a]Pooled analysis was not calculated.

2023382878

ploratory analysis. The strength of the association between cooking fuel and illness was only modestly sensitive to the definitions of the variables (i.e., SES and cooking fuel), and to the number of subjects and city-cohorts. The initial analysis based on 10 cohorts was significant (analysis 1). The reanalysis with modified independent variables and additional cohorts was also significant (analysis 4). A lesser association was found (analysis 2) when parental education was included specifically in the analysis. This effect of model socioeconomic status (parameterized by parental education) is consistent with results reported from a similar study of children from western Pennsylvania (3).

The second example of the sensitivity of the data to the style of analysis concerns our examination of the illness rates in boys and girls.

This is shown in the evaluation of illness rates for four respiratory diseases: asthma, bronchitis, illness before age 2 and illness last year. The data were from 16 city cohorts (except for illness before age 2 which was based on 14 cohorts). Analyses were conducted separately for each sex, thus providing a total of eight analyses (four diseases × two sexes). The risk factors considered in each analysis were: city, cohort, maternal smoking, age, parental education, and cooking fuel. Log-linear analysis was used to study the relationships between each disease and the risk factors. Both step-up and step-down analyses were performed, and where the results diverged, the simpler model was chosen. Details of this analysis are presented elsewhere (4).

Figures 9-12 present the standardized illness rates by city separately for boys and girls. The rates for each disease and for each sex were adjusted by those risk factors shown to be significant ($p < 0.10$) in the analyses. For example, the rates were adjusted for maternal smoking in all analyses except those for asthma in both sexes and for illness last year for boys. The boys' bronchitis rate was adjusted by cooking fuel, parental education and maternal smoking. For bronchitis in girls, the rates were adjusted only by maternal smoking. For illness before age 2 in boys, the rates were adjusted for age as well as for maternal smoking. For illness before age 2 in girls, the rates were adjusted by parental education and maternal smoking. For illness last year in boys,

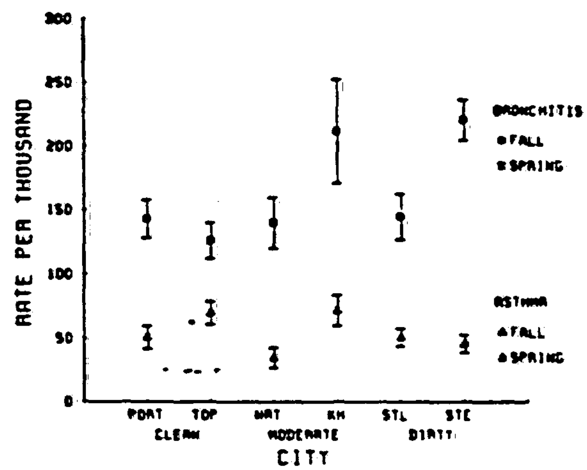

FIGURE 10. Illness rates for bronchitis and asthma for girls by city and season.

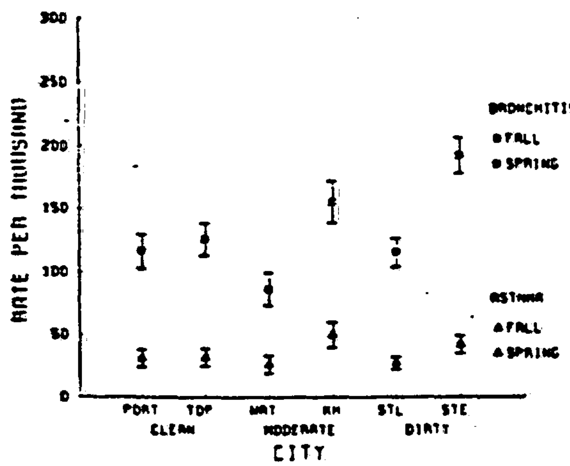

FIGURE 9. Illness rates for bronchitis and asthma for boys by city and season.

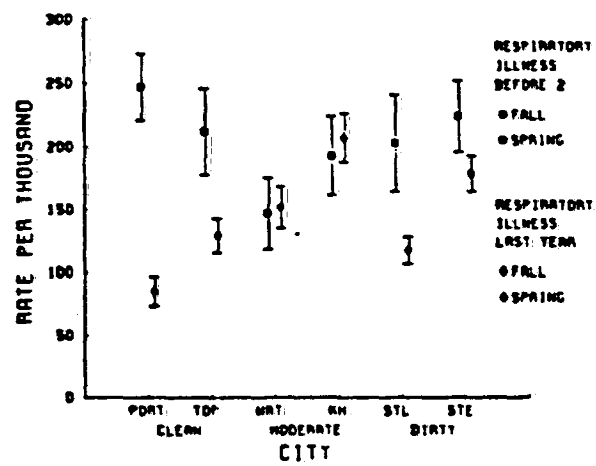

FIGURE 11. Illness rates for respiratory illness before age 2 and illness last year for boys by city and season.

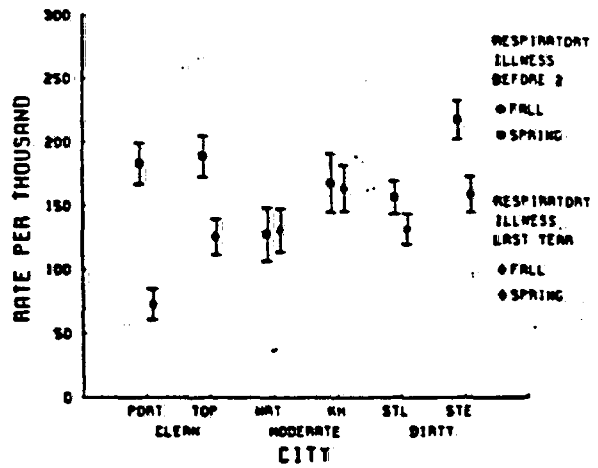

FIGURE 12. Illness rates for respiratory illness before age 2 and illness last year for girls by city and season.

the rates were adjusted by age. For illness last year in girls, the rates were adjusted by age and maternal smoking. The rates for asthma and bronchitis are shown in Figure 9 for boys and in Figure 10 for girls. The rates for illness last year and for illness before age 2 are shown in Figure 11 for boys and in Figure 12 for girls. Although differences between cities for reported history of illness before age 2 varied according to year of enrollment for boys, the rates were averaged over city to simplify the visual presentation.

For all diseases except asthma in girls, city was a significant factor. For asthma in boys, the city with the highest rate was Kingston-Harriman. For bronchitis, the cities with the highest rates for both boys and girls were Steubenville and Kingston-Harriman. For illness before age 2, the highest rate for boys was in Portage, while for girls it was in Steubenville. For illness last year, the highest rates for both sexes were in Kingston-Harriman. Generalizing across the seven analyses, Steubenville and Kingston-Harriman had the highest illness rates and Watertown the lowest.

Three cities (Portage, Watertown and St. Louis) were visited in the fall, and the others were visited during the spring. Because respiratory diseases are more common during the winter, and because a recent disease is more likely to be remembered than one a season or more earlier, it is revealing to examine the rates for the cities grouped by season. With this in mind, Figure 9 shows that bronchitis was higher for boys in the spring cities, and asthma shows little seasonal effect. Figure 10 shows that bronchitis was also higher for girls in the spring cities. Figures 11 and 12 show that illness last year was higher in the spring cities. No trend was evident for illness before age 2. These data suggest that there is a bias in recall by the parents, with greater numbers of events reported for those reporting closest to the winter season just passed.

Maternal smoking was associated with history of bronchitis, illness before age 2 and illness last year in girls (p ranged from 0.04 to 0.06). Except for illness in the last year the same associations were found for maternal smoking and disease in boys (p ranged from 0.01 to 0.08). In all cases, the illness rates were higher for children whose mothers smoked.

Cooking fuel was related only to boys' bronchitis. Boys in homes where gas was used for cooking had lower bronchitis rates than boys in homes using other cooking fuels. This is in contrast to the earlier reported finding concerning illness before age 2 in which children in homes with gas cooking had higher rates than children in homes with other cooking fuels.

Discussion

These examples of cross-sectional analyses illustrate potential pitfalls in analyzing studies of possible health effects of air pollution exposures close to the present ambient air quality standards. The analyses are sensitive to the sample composition and to the assumptions and variable definitions used. In addition, there are many confounding factors and the signal-to-noise ratio is small, so that even in relatively large studies, resolution may be difficult. It is useful to define variables and analysis plans prior to beginning the analysis to avoid subjectivity and data-dependent results. We hope that the prospective aspects of our study will help provide less ambiguous measures of the health significance of low level exposure to fossil fuel air pollutants.

This study was supported by grants from the National Institute of Environmental Health Sciences (ES01108), the Environmental Protection Agency (EP 68-02-3201, 3644), and the Electric Power Research Institute (RP1001-1).

REFERENCES

1. Ferris, B. G., Jr., Speizer, F. E., Spengler, J. D., Dockery, D., Bishop, Y. M. M., Wolfson, M., and Humble, C. Effects of sulfur oxides and respirable particles on human health. Methodology and demography of populations in study. Am. Rev. Resp. Dis. 120: 767-779 (1979).
2. Speizer, F. E., Ferris, B., Jr., Bishop, Y. M. M., and Spengler, J. Respiratory disease rates and pulmonary func-

2023382880

tion in children associated with NO_2 exposure. Am. Rev. Resp. Dis. 121: 3-10 (1980).

Schenker, M. B., Speizer, F. E., and Samet, J. Biologic and environmental determinants of childhood respiratory disease. Paper presented at Annual Meeting, American Thoracic Society, Detroit, May 9-13, 1981.

4. Ware, J. W., Weener, J., Spiro, A., III, Dockery, D. W., Speizer, F. E., Spengler, J. D., and Ferris, B. G., Jr. Cross-sectional analyses of illness history and symptom prevalence in school children in six cities with different air pollution levels. In preparation.